Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 85-39	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On detection of number of signals in presence of colored noise using information theoretic criteria.		5. TYPE OF REPORT & PERIOD COVERED Technical - October 1985
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai		8. CONTRACT OR GRANT NUMBER(s) N00014-85-K-0292 (Navy) F49620-85-C-0008 (Air Force)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis 515 Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research and Air Force Office of Scientific Research		12. REPORT DATE October 1985
		13. NUMBER OF PAGES 33
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Colored noise, information theoretic criteria, likelihood ratio tests, signal detection, and strong consistency.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, the authors proposed information theoretic criteria for detection of the number of signals when the noise is colored. The strong consistency of these criteria is also established.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ON DETECTION OF NUMBER OF SIGNALS IN PRESENCE OF COLORED NOISE USING INFORMATION THEORETIC CRITERIA

L. C. Zhao

P. R. Krishnaiah

Z. D. Bai

Center for Multivariate Analysis
~~University of~~ Pittsburgh University.

October 1985

Technical Report #85-39

Center for Multivariate Analysis
Fifth Avenue, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

*This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

1. INTRODUCTION

In the area of signal processing, a model that is often used involves expressing the observation vector as the sum of noise vector and vector of linear combinations of the (random) signal vector. The noise vector and signal vectors are usually assumed to be distributed independently as normal with zero mean vectors. When the noise is white, the problem of detection of the number of signals transmitted is related to finding the multiplicity of the smallest eigenvalue of the covariance matrix of the observation vector. So, eigenvalue methods play an important role in signal processing. These methods play a dominant role in the area of multivariate statistical analysis. Some workers (e.g., see Kumaresan and Tufts(1980), Liggett(1973), Schmidt(1979), Tufts, Kirsteins and Kurmaresan(1983), Wax and Kailath(1984)) in signal processing have used the eigenvalue methods.

Recently, eigenvalue methods involving information theoretic criteria are used by Wax and Kailath (1984) and Zhao, Krishnaiah and Bai (1985) for determination of the number of signals in presence of white noise. The object of this paper is to detect the number of signals present in presence of colored noise. This problem is equivalent to the problem of studying the rank of Γ when $\Sigma_2 = \Gamma + \lambda\Sigma_1$, λ is a known or unknown scalar, Σ_2 and Σ_1 are $p\times p$ covariance matrices and Γ is nonnegative definite matrix of unknown rank $q < p$. This problem arises in other areas like one-way multivariate components of variance model and factor analysis. Now, let n_1S_1 and n_2S_2 be distributed independently as central Wishart matrices with n_1 and n_2 degrees of freedom and $E(S_i) = \Sigma_i,(i=1,2)$. Rao (1983) derived the likelihood ratio test (LRT) statistic for the rank of Γ when λ is unknown. He also derived a modified LRT statistic for the rank of Γ when λ is known. The main contribution of our paper is to propose certain information theoretic criteria for detection of the number of signals and establish the property of strong consistency. The paper is organized as follows.

In Section 2 of the paper, we discuss the model considered in the case of colored noise. In Section 3, we discuss the LRT and other test procedures for testing the hypothesis that the last few eigenvalues of $\Sigma_2\Sigma_1^{-1}$ are equal to λ for the cases when λ is known and unknown. We propose certain information theoretic criteria in Section 4 for detection of the number of signals when λ is known as well as when it is unknown and establish the strong consistency of these procedures. Some alternative information theoretic criteria are also mentioned. In Section 5, we discuss the applications of the above results to determine the rank of the covariance matrix of random effects vector under multivariate one-way classification model. The results in Sections 3-5 are discussed when the first q eigenvalues $\lambda_1,\ldots,\lambda_q$ of $\Sigma_2\Sigma_1^{-1}$ are simple (distinct). When these eigenvalues have multiplicities (that is, they are equal in groups), the situation becomes complicated. In this case, the problem invlolves not only estimation of q but also the multiplicities of the first q eigenvalues. This problem is investigated in Section 6.

2. A MODEL IN SIGNAL PROCESSING

In the area of signal processing, the following model is used:

$$\underset{\sim}{x}(t) = A\underset{\sim}{s}(t) + \lambda\underset{\sim}{n}(t) \tag{2.1}$$

where $\underset{\sim}{x}(t)$: $p\times 1$ is the observation vector, $A = [A(\underset{\sim}{\Phi}_1),\ldots,A(\underset{\sim}{\Phi}_q)]$, $\underset{\sim}{s}'(t)=(s_1(t)$ $\ldots,s_q(t))$, $s_i(t)$ is a complex waveform which is referred to as i-th signal, $A(\underset{\sim}{\Phi}_i)$ is $p\times 1$ complex vector which depends upon the vector $\underset{\sim}{\Phi}_i$ of unknown parameters associated with i-th signal, $\underset{\sim}{n}(t)$ is a complex vector associated with the noise and λ is known or unknown scalar. We assume that $\underset{\sim}{s}(t)$ and $\underset{\sim}{n}(t)$ are distributed independently as complex multivariate normal with covariance matrices ψ and Σ_1 respectively, $E(\underset{\sim}{s}(t))= \underset{\sim}{0}$, and $E(\underset{\sim}{n}(t)) = \underset{\sim}{0}$. Also, $\bar{A}$ denotes the complex conjugate of A and A^* denotes the transpose of $\bar{A}$. The number of signals, q, transmitted is equal to the rank of $A\Psi A^*$. If $\lambda_1 \geq \ldots \geq \lambda_p$ denote the eigenvalues of $\Sigma_2\Sigma_1^{-1}$, q is given by

$$\lambda_1 \geq \cdots \geq \lambda_q > \lambda_{q+1} = \cdots = \lambda_p = \lambda \tag{2.1}$$

since $\Sigma_2 = A\Psi A^* + \lambda\Sigma_1$. When $\Sigma_1 = \sigma^2 I$, the problem of determination of the number of signals was considered in the literature. Wax and Kailath (1985) used Akaike's AIC criterion and the minimum discription length (MDL) criterion due to Rissanen and Schwartz when σ^2 is unknown, $\lambda=1$ and the underlying distribution is complex multivariate normal. Zhao, Krishnaiah and Bai (1985) considered an alternative criterion and established its strong consistency for the cases when σ^2 is known and unknown when the underlying distribution is complex multivariate normal. They have also considered certain cases when the underlying distribution is not necessarily complex multivariate normal. But, it is not realistic to assume always that the noise is white. We assume that the covariance matrix Σ_1 of $\underset{\sim}{n}(t)$ is arbitrary and an independent estimate S_1 of Σ_1 is available from a different data set. Also, we assume that n independent observations $\underset{\sim}{x}(t_1),\ldots,\underset{\sim}{x}(t_n)$ are available on $\underset{\sim}{x}(t)$. In this case, we can estimate Σ_2 with S_2 where $n_2S_2 = \sum_{j=1}^{n_2} \underset{\sim}{x}(t_j)\underset{\sim}{x}^*(t_j)$. Since S_1 and S_2 are distributed independently as complex Wishart matrices with n_1 and n_2 degrees of freedom respectively, $E(S_1) = \Sigma_1$ and $E(S_2) = \Sigma_2 = A\Psi A^* + \Sigma_1$, the methods developed in this paper are useful in finding the number of signals transmitted.

We will develop the methodology for finding q such that $\lambda_1 \geq \cdots \geq \lambda_q > \lambda_{q+1} = \cdots = \lambda_p = \lambda$ where $\lambda_1 \geq \cdots \geq \lambda_p$ are the eigenvalues of $\Sigma_2\Sigma_1^{-1}$, n_1S_1 and n_2S_2 are distributed independently as real central Wishart matrices with n_1 and n_2 degrees of freedom, $E(S_i) = \Sigma_i (i=1,2)$, and $\Sigma_2 = A\psi A' + \lambda\Sigma_1$ when λ is a real scalar. Here A: p×q is a real unknown matrix, Ψ: q×q is a real, positive definite matrix. The above methodology needs only trivial modification when n_1S_1 and n_2S_2 are complex Wishart matirces.

3. TESTS FOR THE EQUALITY OF THE LAST FEW EIGENVALUES OF $\Sigma_2\Sigma_1^{-1}$

Let n_1S_1 and n_2S_2 be distributed independently as central Wishart matrices with n_1 and n_2 degrees of freedom respectively, $E(S_1) = \Sigma_1$, $E(S_2) = \Sigma_2$ and $\Sigma_2 = A\psi A' + \lambda\Sigma_1$. The log likelihood function $L(\theta)$ is given by

$$2L(\theta) = - n_1\log|\Sigma_1| - n_2\log|\Sigma_2| - n_1\mathrm{tr}\Sigma_1^{-1}S_1 - n_2\mathrm{tr}\Sigma_2^{-1}S_2. \tag{3.1}$$

Let H_k: $\lambda_1 \geq \cdots \geq \lambda_k > \lambda_{k+1} = \cdots = \lambda_p = 1$. We first calculate $\sup_{\theta\in\Theta_k} L(\theta)$ where Θ_k is the paramteric space under H_k. Let the eigenvalues of $S_2S_1^{-1}$ be $\delta_1 \geq \cdots \geq \delta_p$. With probability one we have $\delta_1 > \delta_2 > \cdots > \delta_p > 0$. We know that there exists two nonsingular matrices R and $\hat{R}$ such that

$$\begin{aligned} \Sigma_1 &= RR', & \Sigma_2 &= R\Lambda R' \\ S_1 &= \hat{R}\hat{R}', & S_2 &= \hat{R}\Delta\hat{R}', \end{aligned} \tag{3.2}$$

where $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ and $\Delta = \mathrm{diag}(\delta_1,\ldots,\delta_p)$. Let $R^{-1}\hat{R} = V$. Then

$$2L(\theta) = - (n_1+n_2)\log|\hat{R}\hat{R}'| - n_2\log(\lambda_1\cdots\lambda_p) + L_1(V',\Lambda), \tag{3.3}$$

where

$$L_1(V',\Lambda) = (n_1+n_2)\log|V'V| - n_1\mathrm{tr}V'V - n_2\mathrm{tr}(\Lambda^{-1}V\Delta V'). \tag{3.4}$$

First we fix $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ and compute $\mathrm{Sup}_{V'}L_1(V',\Lambda)$. If we take partial derivative of L with respect to V', we obtain the following optimizing equations:

$$2(n_1+n_2)V^{-1} - 2n_1V' - 2n_2\Delta V'\Lambda^{-1} = 0$$

i.e.,

$$\alpha_n V'V + \beta_n \Delta V'\Lambda^{-1}V = I_p, \tag{3.5}$$

where I_p is the p×p identity matrix, and

$$\alpha_n = n_1/n, \quad \beta_n = n_2/n, \quad n = n_1 + n_2. \tag{3.6}$$

From (3.5) it follows that $\Delta V'\Lambda^{-1}V$ is symmetric and hence $\Delta V'\Lambda^{-1}V = V'\Lambda^{-1}V\Delta$. Since $\delta_1 > \delta_2 > \ldots > \delta_p > 0$, $\Lambda V'\Lambda^{-1}V$ is diagonal so that by (3.5) $V'V$ is diagonal. Thus there is an orthogonal matrix Q and a diagonal matrix $D = \text{diag}[d_1,\ldots,d_p]$, $d_1 \geq d_2 \geq \ldots \geq d_p > 0$ such that $V = QD$. Since Δ and $\Delta V'\Lambda^{-1}V$ are diagonal, $\Delta V'\Lambda^{-1}V = \Delta^{\frac{1}{2}}V'\Lambda^{-1}V\Delta^{\frac{1}{2}} = \Delta^{\frac{1}{2}}DQ'\Lambda^{-1}QD\Lambda^{\frac{1}{2}}$ so that $Q'\Lambda^{-1}Q$ is also diagonal and the diagonal elements are the same as those of Λ^{-1}. Again by (3.5) we know that the diagonal elements are arranged according to the increasing order. Hence $Q'\Lambda^{-1}Q = \Lambda^{-1}$ or equivalently,

$$\Lambda^{-1}Q = Q\Lambda^{-1} \tag{3.7}$$

Substituting this into (3.5) we find

$$\begin{aligned} I_p &= \alpha_n V'V + \beta_n \Delta V'\Lambda^{-1}V = \alpha_n V'V + \beta_n V'\Lambda^{-1}V\Delta \\ &= \alpha_n V'V + \beta_n DQ'\Lambda^{-1}QD\Lambda = \alpha_n V'V + \beta_n DQ'QD\Lambda^{-1}\Lambda \\ &= V'V(\alpha_n I + \beta_n \Lambda^{-1}\Delta) \end{aligned} \tag{3.8}$$

and

$$|V'V| = |\alpha_n I + \beta_n \Lambda^{-1}\Lambda|^{-1} = \prod_{i=1}^{p} \frac{1}{(\alpha_n + \beta_n \lambda_i^{-1}\delta_i)} . \tag{3.9}$$

Also, we have by (3.5)

$$- n_1 \text{tr}V'V - n_2 \text{tr}(\Lambda V'\Lambda^{-1}V) = - (n_1+n_2)p. \tag{3.10}$$

By (3.4), (3.9) and (3.10)

$$\sup_{V'} L_1(V',\Lambda) = (n_1+n_2)\{-\sum_{i=1}^{p} \log(\alpha_n+\beta_n\lambda_i^{-1}\delta_i) - p\} \tag{3.11}$$

So,

$$\sup_{\theta\in\Theta_k} 2L(\theta) = \sup_{\lambda_1\geq\cdots\geq\lambda_k>1} n\{-\log|\hat{R}\hat{R}'| - \beta_n\log(\lambda_1\cdots\lambda_k) - p - \sum_{i=k+1}^{p}\log(\alpha_n+\beta_n\delta_i) - \sum_{i=1}^{k}\log(\alpha_n+\beta_n\lambda_i^{-1}\delta_i)\}, \tag{3.12}$$

where Θ_k denotes the parametric space when H_k is true. Let $\tau = \#\{i \leq p: \delta_i > 1\}$. Also, let $d = \min\{k,\tau\}$, and set

$$\phi_1 = n\sum_{i=1}^{d}\{-\log(\alpha_n+\beta_n\lambda_i^{-1}\delta_i) - \beta_n\log\lambda_i\}$$
$$\phi_2 = n\sum_{i=d+1}^{k}\{-\log(\alpha_n+\beta_n\lambda_i^{-1}\delta_i) - \beta_n\log\lambda_i\}. \tag{3.13}$$

We note that $\delta_1 > \delta_2 > \ldots > \delta_d > 1$, and $\sup_{\lambda_1>\cdots>\lambda_d>1}\phi_1$ can be reached at $\lambda_i = \delta_i$, for $i=1,\ldots,d$. For $i=d+1,\ldots,k$, $\delta_i < 1$ and $\lambda_i > 1$, we see that the function

$$f_i(\lambda_i) = -\log(\alpha_n+\beta_n\lambda_i^{-1}\delta_i) - \beta_n\log\lambda_i \tag{3.14}$$

has negative derivative, and $f_i(\lambda_i)$ is decreasing and continuous. Thus

$$\sup_{\lambda_i>1} f_i(\lambda_i) = f_i(1), \quad i=d+1,\ldots,k.$$

From the above discussion, we have

$$\sup_{\lambda_1>\cdots>\lambda_d>1}\phi_1 = -n\beta_n\sum_{i=1}^{d}\log\delta_i, \tag{3.15}$$

and

$$\sup_{\lambda_{d+1}>\cdots>\lambda_k>1}\phi_2 = -n\sum_{i=d+1}^{k}\log(\alpha_n+\beta_n\delta_i). \tag{3.16}$$

From (3.12),(3.13)-(3.15), it follows that

$$\begin{aligned}\sup_{\theta\in\Theta_k} 2L(\theta) &= -n\{\log|\hat{R}\hat{R}'| + p + \beta_n \sum_{i=1}^{p} \log\delta_i\} \\ &\quad - n\sum_{i=1+\min(k,\tau)}^{p} [\log(\alpha_n+\beta_n\delta_i) - \beta_n\log\delta_i] \\ &= -n_1\log|S_1| - n_2\log|S_2| - np \qquad (3.17)\\ &\quad - n\sum_{i=1+\min(k,\tau)} [\log(\alpha_n+\beta_n\delta_i) - \beta_n\log\delta_i] .\end{aligned}$$

So, the LRT statistic for testing the hypothesis H_k against the alternative that the rank is more than k is given by

$$L_k = \prod_{i=1+\min(k,\tau)}^{p} (\alpha_n+\beta_n\delta_i)^{-n/2}\delta_i^{n\beta_n/2} \qquad (3.18)$$

The LRT statistic for testing H_k against H_t $(k<t)$ is given by

$$L_{kt} = \prod_{i=1+\min(k,\tau)}^{\min(t,\tau)} (\alpha_n+\beta_n\delta_i)^{-n/2}\delta_i^{n\beta_n/2} \qquad (3.19)$$

if $k < \tau$. If $k \geq \tau$, then $L_{kt} = 1$.

Rao (1983) considered the problem of testing the hypothesis that the rank of Γ is k against the alternative that it is greater than k when $\Sigma_2 = \Gamma + \lambda\Sigma_1$ and Γ is nonnegative definite for the cases when $\lambda = 1$ and when λ is unknown. He proposed a modified LRT procedure and the LRT procedure for testing the hypothesis on the rank of Γ according as $\lambda = 1$ and λ is unknown.

When λ is unknown, let H_k^* denote the hypothesis that

$$\lambda_1 \geq \cdots \geq \lambda_k > \lambda_{k+1} = \cdots = \lambda_p = \lambda$$

for $k = 0,1,\ldots,(p-1)$. Let M_k^* denote the model for which H_k^* is true. It is known (see Rao(1983)) that the supremum of the logarithm of the likelihood function under H_k^* is given by

$$- \frac{n_1}{2} \log|S_1| - \frac{n_2}{2} \log|S_2| - \frac{np}{2}$$
$$+ \frac{n}{2} \sum_{i=k+1}^{p} [\alpha_n \log\hat{\lambda}_{ko} + \beta_n \log\delta_j - \log(\alpha_n\hat{\lambda}_{ko} + \beta_n\delta_j)] \tag{3.20}$$

where $\hat{\lambda}_{k0}$ satisfies the equation

$$p-k = \sum_{j=k+1}^{p} \frac{\delta_j}{\alpha_n\hat{\lambda}_{ko}+\beta_n\delta_j} \tag{3.21}$$

or equivalently,

$$p-k = \sum_{j=k+1}^{p} \frac{\hat{\lambda}_{ko}}{\alpha_n\hat{\lambda}_{ko}+\beta_n\delta_j}. \tag{3.22}$$

As pointed out in Rao (1983), the logarithm of the likelihood ratio statistic is given by

$$- 2\log L_k^* = \log \prod_{i=k+1}^{p} [((n_2\delta_i+n_1\hat{\lambda}_{ko})/n)^n \frac{1}{\delta_i^{n_2}\hat{\lambda}_{ko}^{n_1}} \tag{3.23}$$

which is distributed as chi-square with $[(p-k)(p-k+1)-2]/2$ degrees of freedom as n_1 and n_2 tend to infinity.

We will propose the following alternative procedures for testing the hypothesis H_k against the alternative that $\lambda_{k+1} > 1$. We accept or reject H_k against $\lambda_{k+1} > 1$ according as

$$g(\ell_{k+1},\ldots,\ell_p) \lessgtr c_\alpha \tag{3.24}$$

where

$$P[g(\ell_{k+1},\ldots,\ell_p) \le c_\alpha | H_k] = (1-\alpha). \tag{3.25}$$

For example, $g(\ell_{k+1},\ldots,\ell_p)$ may be ℓ_{k+1} or $\ell_{k+1}+ \ldots + \ell_p$. The exact distributions of the above statistics are complicated. Also, they involve nuisance parameters unless k = 0. But, the joint asymptotic distribution of $\ell_{k+1},\ldots,\ell_p$, is given in a companion paper (in preparation) by the authors for the real and complex cases. A review of other asymptotic results was given in Murihead (1978). We can use the above result to obtain asymptotic distributions of statistics like ℓ_{k+1} and $\ell_{k+1} + \ldots + \ell_p$.

We will now consider the case when λ is unknown. Let H_{ij}: $\lambda_i = \lambda_j$. Then H_k^* can be decomposed as $\bigcap_{i=k+1}^{p-1} H_{ip}$, $\bigcap_{i=k+1}^{p-1} H_{i,i+1}$ and $\bigcap_{i<j=k+1}^{p} H_{ij}$. Motivated by the above decompositions, we propose the following procedures. We accept H_k^* against $\bigcup_{i=k+1}^{p-1} [\lambda_i > \lambda_p]$ if

$$(\ell_i/\ell_p) \le c_{\alpha 1} \tag{3.26}$$

for i = k+1,...,p-1 and reject it otherwise where

$$P[(\ell_{k+1}/\ell_p) \le c_{\alpha 1} | H_k^*] = (1-\alpha). \tag{3.27}$$

If H_k^* is rejected, we accept or reject the subhypothesis H_{k+1}^* according as

$$\ell_i/\ell_p \lessgtr c_{\alpha 1} \tag{3.28}$$

for i=k+2,...,p-1.

The hypothesis H_k^* when tested against $\bigcup_{i=k+1}^{p-1} (\lambda_i > \lambda_{i+1})$ is accepted if

$$(\ell_i/\ell_{i+1}) \leq c_{\alpha 2} \tag{3.29}$$

for $i = k+1,2,\ldots,(p-1)$ and rejected otherwise where

$$P[(\ell_i/\ell_{i+1}) \leq c_{\alpha 2};\ i = k+1,2,\ldots,p-1 | H_k^*] = (1-\alpha). \tag{3.30}$$

Similarly, the hypothesis H_k^* when tested against $\bigcup_{i<j=k+1}^{p} [\lambda_i > \lambda_{i+1}]$ is accepted if and only if

$$\ell_i/\ell_j \leq c_{\alpha 3}$$

for $i < j$ and $i = k+1,\ldots,p-1$ where

$$P[(\ell_{k+1}/\ell_p) \leq c_{\alpha 3} | H_k^*] = (1-\alpha). \tag{3.31}$$

We can also use $\ell_{k+1}/(\ell_{k+1}+\ldots+\ell_p)$ as a test statistic. Exact computations of the probability integrals associated with the above procedures are complicated and involve nuisance parameters except when $k = 0$. In this particular case, percentage points are available for a few special cases (see Krishnaiah (1980)). But, approximations to the critical values $c_{\alpha 1}$, $c_{\alpha 2}$ and $c_{\alpha 3}$ can be derived for large samples.

4. DETECTION OF NUMBER OF SIGNALS USING INFORMATION THEORETIC CRITERIA

In the preceding section, we discussed various procedures for testing the hypotheses on the number of signals. We will now discuss procedures for detection of the number of signals by using information theoretic criteria.

When $\lambda = 1$, the likelihood ratio test statistic L_k for H_k is given by (3.18). Now, let

$$G(k) = \log L_k$$

and

$$I(k,C_n) = -\log L_k + C_n \nu(k,p) \tag{4.1}$$

where $\nu(k,p) = \frac{1}{2}k(2p-k+1)$ denotes the number of free parameters and C_n satisfies the following condition:

$$\text{(i)} \quad \lim_{n\to\infty}(C_n/n) = 0 \tag{4.2}$$

$$\text{(ii)} \quad \lim_{n\to\infty}(C_n/\log\log n) = \infty. \tag{4.3}$$

Then, according to FDC criterion we find $\hat{q}$ which satisfies

$$I(\hat{q},C_n) = \min\{I(0,C_n),\ldots,I(p-1,C_n)\} \tag{4.4}$$

and use $\hat{q}$ as an estimate of q which is the number of signals present in the true model M_q. The strong consistency of $\hat{q}$ is proved below:

<u>THEOREM 4.1</u>. If $S_i \sim W_p(n_i,\Sigma_i)$, $i=1,2$, $n = n_1+n_2 \to \infty$ and $\alpha_n \in [a,b] \subset (0,1)$ with a,b being constants, then $\hat{q}$ is a strongly consistent estimate of q under the model M_q.

<u>PROOF</u>. It is known (see Zhao, Krishnaiah and Bai (1985)) that

$$|\delta_i - \lambda_i| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.} \tag{4.5}$$

Using Taylor's expansion, we get

$$\log(\alpha_n+\beta_n\delta_i) - \beta_n\log\delta_i = \frac{1}{2}\alpha_n\beta_n(\delta_i-1)^2(1+o(1)) \quad \text{a.s.} \tag{4.6}$$

for $i > q$. Here we used the fact that

$$\lim_{n\to\infty} \delta_i = 1 \quad \text{a.s.} \quad \text{for} \quad i > q. \tag{4.7}$$

With probability one for large n, we have $\min(q,\tau) = q$ so that, for large n,

$$|G(q) - G(k)| = O(\log\log n) = o(C_n) \tag{4.8}$$

and

$$\begin{aligned}(I(q,C_n) - I(k,C_n))/C_n &= -(G(q) - G(k)/C_n - (k-q)(2p-k-q+1) \\ &\to -(k-q)(2p-k-q+1)\end{aligned} \tag{4.9}$$

when $k > q$. Thus with probability one, for large n,

$$I(q,C_n) < I(k,C_n), \quad \text{if } k > q. \tag{4.10}$$

Since, with probability one, $\min(q,\tau) = q$, for large n, we have

$$G(q) - G(k) = n\sum_{i=k+1}^{q}[\log(\alpha_n+\beta_n\delta_i) - \beta_n\log\delta_i] \tag{4.11}$$

for $k < q$. Note that for $i \leq q$, $\lim_{n\to\infty} \delta_i = \lambda_i > 1$, we see that there exists a constant $\mu > 1$ such that with probability one

$$\delta_i > \mu \text{ for } i = k+1,\ldots,q,$$

for large n. By the monotonicity of

$$f(\delta) = \log(\alpha_n+\beta_n\delta) - \beta_n\log\delta, \tag{4.12}$$

$$P(CD) = P[I(q,C_n) - I(k,C_n) < 0;\ k=0,1,\ldots,(p-1);\ k \neq q | H_q]. \tag{4.20}$$

It would be of interest to compare numerically the probability of correct detection of the criteria $I(\hat{q},C_N)$, AIC($\tilde{q}$) and MDL($\bar{q}$).

We now discuss the problem of detection of the number of signals when λ is unknown. In this case, the logarithm of the LRT statistic for H_k^* is given by (3.23). Now, let

$$G^*(k) = \log L_k^* \tag{4.21}$$

and assume that C_n satisfies the conditions (4.2) and (4.3). Then an estiamte of q, the true number of signals, is given by $\hat{q}$ where

$$\hat{q} = \max\{k:\ 1 \leq k \leq p-1,\ G^*(k) - G^*(k-1) > C_n\} \tag{4.22}$$

and $\max\phi = 0$ for convenience.
Let M_k^* denote the model under which H_k^* is true. We now prove the strong consistency of the above method.

THEOREM 4.2. If $S_i \sim W_p(n_i,\Sigma_i)$, $i = 1,2$, $n \to \infty$, and $\alpha_n \in [a,b] \subset (0,1)$ with a,b being constants, then, under the true model M_q^*, $\hat{q}$ is a strongly consistent estimate of q.

PROOF. As pointed out earlier

$$|\delta_j - \lambda_j| = O\left(\sqrt{\frac{1}{n}\log\log n}\right) \text{ a.s.} \tag{4.23}$$

for $j = 1,\ldots,p$.

Note that being $\delta_1 > \delta_2 > \ldots > \delta_p$ with probability one, we can see from (3.21) or (3.22) that $\hat{\lambda}_{ko} > \hat{\lambda}_{k+1,0}$ for $0 \leq k \leq p-1$.

We assume that M_q^* is the true model, and $k \geq q$. As mentioned above,

$$|\delta_j - 1| = O(\sqrt{\tfrac{1}{n}\log\log n}) \quad \text{a.s.}$$

for $j \geq q$. Assume $|\delta_j - 1| \geq \varepsilon_n$ for $j \geq k+1$. Then

$$\frac{p-k}{\alpha_n\hat{\lambda}_{k0}(1-\varepsilon_n)^{-1}+\beta_n} \leq \sum_{j=k+1}^{p} \frac{\delta_j}{\alpha_n\hat{\lambda}_{k0}+\beta_n\delta_j} = p-k \leq \frac{(p-k)}{\alpha_n\hat{\lambda}_{k0}(1+\varepsilon_n)^{-1}+\beta_n}, \tag{4.24}$$

and $1 - \varepsilon_n \leq \hat{\lambda}_{k0} \leq 1 + \varepsilon_n$. Thus we have

$$|\hat{\lambda}_{k0} - 1| = O(\sqrt{\tfrac{1}{n}\log\log n}) \quad \text{a.s.}, \ k \geq q. \tag{4.25}$$

Using Taylor's expansion, we see that for $k \geq q$,

$$n\sum_{j=k+1}^{p}[\alpha_n\log\hat{\lambda}_{k0} + \beta_n\log\delta_j - \log(\alpha_n\hat{\lambda}_{k0}+\beta_n\delta_j)]$$

$$\overset{\text{a.s.}}{=} \frac{n}{2}\sum_{j=k+1}^{p}[-\alpha_n(\hat{\lambda}_{k0}-1)^2 - \beta_n(\delta_j-1)^2 + (\alpha_n(\hat{\lambda}_{k0}-1) + \beta_n(\delta_j-1))^2](1+o(1))$$

$$= O(\log\log n) \quad \text{a.s.}$$

If $k > q$, then from $\lim_{n\to\infty} C_n/\log\log n = \infty$, we get

$$L^*(k) - L^*(k-1) = o(C_n) \quad \text{a.s.} \tag{4.26}$$

Thus, with probability one, we have for large n,

$$L^*(k) - L^*(k-1) < C_n \quad \text{for} \quad k > q, \tag{4.27}$$

which implies $\hat{q} = 0$ if $q = 0$. Now we assume that $1 \leq k < q$. We have

$$G^*(k) - G^*(k-1) = n\{g(\hat{\lambda}_{k0}) - g(\hat{\lambda}_{k-1,0}) - \alpha_n\log\hat{\lambda}_{k-1,0} - \beta_n\log\delta_k$$

$$+ \log(\alpha_n\hat{\lambda}_{k-1,0}+\beta_n\delta_k)\}, \tag{4.28}$$

where

$$g(x) = \alpha_n(p-k)\log x - \sum_{j=k+1}^{p} \log(\alpha_n x+\beta_n\delta_i) \tag{4.29}$$

It is easily seen that for $x\epsilon(\hat{\lambda}_{k0},\hat{\lambda}_{k-1,0}]$,

$$g'(x) = \frac{\beta_n}{x}\left[\sum_{j=k+1}^{p} \frac{\delta_j}{\alpha_n x+\beta_n\delta_j} - (p-k)\right]< 0,$$

So we have

$$g(\hat{\lambda}_{k0}) - g(\hat{\lambda}_{k-1,0}) > 0. \tag{4.30}$$

Let μ_n be the solution of the following equation:

$$p - k + 1 = \sum_{j=k}^{p} \frac{\lambda_j}{\alpha_n\mu_n+\beta_n\delta_j}. \tag{4.31}$$

From

$$p - k + 1 = \sum_{j=k}^{p} \frac{\delta_j}{\alpha_n\hat{\lambda}_{k-1,0}+\beta_n\delta_j}$$

and $\lim_{n\to\infty} \delta_j = \lambda_j$ a.s., it follows that

$$\lim_{n\to\infty}\left(\hat{\lambda}_{k-1,0}-\mu_n\right) = 0 \quad \text{a.s.} \tag{4.32}$$

Thus we have

$$n\{\log(\alpha_n\hat{\lambda}_{k-1,0}+\beta_n\delta_k) - \alpha_n\log\hat{\lambda}_{k-1,0} - \beta_n\log\delta_k\}$$

$$\underline{\underline{\text{a.s.}}}\; n\{\log(\alpha_n\mu_n+\beta_n\lambda_k) - \alpha_n\log\mu_n - \beta_n\log\lambda_k\}\{1 + o(1)\} \tag{4.33}$$

Since

$$\sum_{j=k}^{p} \frac{\lambda_j}{\alpha_n\lambda_k+\beta_n\lambda_j} = 1 + \sum_{j=k+1}^{p} \frac{\lambda_j}{\alpha_n\lambda_k+\beta_n\lambda_j} < 1 + p - k,$$

it follows $\mu_n < \mu_o < \lambda_k$ for some constant μ_o.

Hence

$$\begin{aligned} &\log(\alpha_n\mu_n+\beta_n\lambda_k) - \alpha_n\log\mu_n - \beta_n\log\lambda_k \geq \\ &\log(\alpha_n\mu_o+\beta_n\lambda_k) - \alpha_n\log\mu_o - \beta_n\log\lambda_k \overset{\Delta}{=} h(\alpha). \end{aligned} \tag{4.34}$$

Note that $h(\alpha)$ is continuous for $\alpha\epsilon[a,b]$ where $\beta = 1-\alpha$. So there exists a $\alpha_o\epsilon[a,b]$ such that

$$h(\alpha_n) \geq h(\alpha_o) = \log(\alpha_o\mu_o+\beta_o\lambda_k) - \alpha_o\log\mu_o - \beta_o\log\lambda_k > 0 \tag{4.35}$$

By (4.33) - (4.35), we see that with probability one for n large,

$$\begin{aligned} &n\{\log(\alpha_n\hat{\lambda}_{k-1,0}+\beta_n\delta_k) - \alpha_n\log\hat{\lambda}_{k-1,0} - \beta_n\log\delta_k\} \\ &> \frac{1}{2}\, nh(\alpha_o). \end{aligned} \tag{4.36}$$

From (4.28),(4.30),(4.36) and $C_n/n \to 0$, it follows that with probability one for n large,

$$L(k) - L(k-1) > MC_n,\ k = 1,2,\ldots,q \tag{4.37}$$

for any fixed $M > 0$.

By (4.27) and (4.37) we see that with probability one for n large,

$$\hat{q} = q, \tag{4.38}$$

and the theorem is proved.

5. MULTIVARIATE ONE-WAY RANDOM EFFECTS MODEL

In this section, we discuss the relationship between drawing inference on the rank of the covariance matrix of column effects in one-way multivariate random effects model and the problem of finding the number of signals discussed in the preceding sections. The one-way multivariate random effects model is given by

$$\underset{\sim}{x}_{ij} = \underset{\sim}{\mu} + \underset{\sim}{\alpha}_i + \underset{\sim}{\epsilon}_{ij} \tag{5.1}$$

for $i = 1,2,\ldots,k$ $j = 1,2,\ldots,m$ where $\underset{\sim}{\mu}$ is the general mean vector, $\underset{\sim}{\alpha}_i$ is the vector of random effects of i-th column and $\underset{\sim}{x}_{ij}$ denotes the j-th observation on i-th column and $\underset{\sim}{\epsilon}_{ij}$ is distributed as multivariate normal with mean vector $\underset{\sim}{0}$ and covariance matrix Σ_1. Also, $\underset{\sim}{\alpha}_i$ is distributed independent of $\underset{\sim}{\epsilon}_{ij}$ as multivariate normal with mean vector $\underset{\sim}{0}$ and covariance matrix ψ. The covariance matrix of $\underset{\sim}{x}_{ij}$ is Σ_2 where $\Sigma_2 = \psi + \Sigma_1$. It is of interest to test the rank of ψ is r. If the rank of ψ is r, then there exists a full rank matrix $B:(p-r)\times p$ such that $B\psi = 0$. Testing the hypothesis that the rank is zero is equivalent to testing the hypothesis of no column effects. If ψ is not of full rank, then we can take advantage of this knowledge in estimating ψ. Anderson (1984,1985) and Schott and Saw (1984) have independently derived the LRT statistic for testing the hypothesis on the rank of ψ. Now let S_b and S_w respectively denote the between groups and within group sums of squares and cross products matrices respectively. Then S_b and S_w are distributed independently as central Wishart matrices with $(k-1)$ and $(km-k)$ degrees of freedom respectively, $E(S_b) = (k-1)(\Sigma_1+m\psi)$ and $E(S_w) = (km-k)\Sigma_1$.

6. DETECTION OF NUMBER OF SIGNALS WHEN EIGENVALUES OF $\Sigma_2\Sigma_1^{-1}$ HAVE MULTIPLICITIES

In Sections 3 and 4, we discussed the problem of detection of the number of signals under the assumption that the nonzero eigenvalues of the matrix $A\Psi\bar{A}'$ are distinct. But situations arise when it is unrealistic to make the above assumption. In this section, we consider the problem of detecting the number of signals and finding the multiplicities of the eigenvalues of $A\Psi\bar{A}'$. We will first discuss the problems of finding the rank of Γ when the underlying distribution is real multivariate normal.

For the interval [o,p], there exists $\sum_{\ell=1}^{p}\binom{p-1}{\ell-1} = 2^{p-1}$ different integer partitions such as $0=k_0 < k_1 < \dots < k_\ell = p$, $\ell = 1,2,\dots,p$. We denote the set of all such partitions with K. Let

$$H^*_{k_1\dots k_\ell}: \lambda_{k_{i-1}+1} = \dots = \lambda_{k_i} = c_i;\ i = 1,2,\dots,\ell \tag{6.1}$$

where $c_1 > \dots > c_\ell$ are unknown constants and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ are eigenvalues of $\Sigma_2\Sigma_1^{-1}$. We will denote the corresponding parametric space and model with $\Theta^*_{k_1\dots k_\ell}$ and $M^*_{k_1\dots k_\ell}$ respectively. We are interested in selecting the correct model $M^*_{q_1,\dots,q_r}$ using information theoretic criteria when we do not have any knowledge of $q_1,\dots,q_r$. When $H^*_{k_1\dots k_\ell}$ is true, the log-likelihood function $L^*(\theta)$ is given by

$$2L^*(\theta) = -n_1\log|\Sigma_1| - n_2\log|\Sigma_2| - n_1\mathrm{tr}\Sigma_1^{-1}S_1 - n_2\mathrm{tr}\Sigma_2^{-1}S_2. \tag{6.2}$$

At first we calculate $\sup_{\theta\in\Theta^*_{k_1,\dots,k_\ell}} 2L^*(\theta)$.

Denote the eigenvalues of $S_2S_1^{-1}$ by $\delta_1 \geq \dots \geq \delta_p$. There exists two non-singular matrices R and $\hat{R}$ such that

$$\Sigma_1 = RR', \quad \Sigma_2 = R\Lambda R',$$

$$S_1 = \hat{R}\hat{R}', \quad S_2 = \hat{R}\Delta\hat{R}',$$

where $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ and $\Delta = \mathrm{diag}(\delta_1,\ldots,\delta_p)$. Without loss of generality, we assume that $\delta_1 > \ldots > \delta_p > 0$. Now, let $R^{-1}\hat{R} = V$. Then

$$2L^*(\theta) = -(n_1+n_2)\log|\hat{R}\hat{R}'| - n_2\log(\lambda_1\ldots\lambda_p) + L_1(V\ ,\Lambda), \qquad (6.3)$$

where

$$L_1(V\ ,\Lambda) = (n_1+n_2)\log|V'V| - n_1 \mathrm{tr} V'V - n_2 \mathrm{tr}(V'\Lambda^{-1}V\Delta). \qquad (6.4)$$

First we fix $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ where (6.1) is satisfied. For given Λ, we can now calculate $\mathrm{Sup}_V\, L_1(V\ ,\Lambda)$. To accomplish this, we take derivative of L_1 with respect to matrix V' and obtain the following optimizing equations:

$$2(n_1+n_2)V^{-1} - 2n_1V' - 2n_2\Delta V'\Lambda^{-1} = 0$$

i.e.,

$$\alpha_n V'V + \beta_n \Delta V'\Lambda^{-1}V = I_p, \qquad (6.5)$$

where

$$\alpha_n = n_1/n, \quad \beta_n = n_2/n, \quad n = n_1 + n_2. \qquad (6.6)$$

Using the same argument as used to prove (3.8), we find

$$V'V(\alpha_n I_p + \beta_n \Lambda^{-1}\Delta) = I_p, \qquad (6.7)$$

and

$$|V'V| = |\alpha_n I_p + \beta_n\Lambda^{-1}\Delta|^{-1} = \prod_{i=1}^{p} \frac{\lambda_i}{\alpha_n\lambda_i + \beta_n\delta_i}\ . \qquad (6.8)$$

Also, we have by (6.5)

$$- n_1 \mathrm{tr} V'V - n_2 \mathrm{tr}(\Delta V' \Lambda^{-1} V) = - (n_1+n_2)p. \tag{6.9}$$

Now, using (6.3), (6.4), (6.8) and (6.9), we get

$$\sup_{\theta\in\Theta^*_{k_1,\ldots,k_\ell}} 2L^*(\theta) = \sup_{c_1>\ldots>c_\ell} n\{- \log|\hat{R}\hat{R}'| - p + \alpha_n \sum_{i=1}^{\ell} \Delta k_i \log c_i - \sum_{i=1}^{\ell} \sum_{j\in\kappa_i} \log(\alpha_n c_i + \beta_n \delta_j)\} \tag{6.10}$$

where $\kappa_i = \{k_{i-1}+1, k_{i-1}+2,\ldots,k_i\}$. Now, let $\hat{c}_i$'s be chosen such that

$$\sum_{j\in\kappa_i} \frac{\delta_j}{\alpha_n \hat{c}_i + \beta_n \delta_j} = \Delta k_i, \; i = 1,\ldots,\ell. \tag{6.11}$$

Noticing that $\hat{c}_i \geq \delta_{k_i} > \delta_{k_i+1} \geq \hat{c}_{i+1}$, then we get

$$\sup_{\theta\in\Theta^*_{k_1,\ldots,k_\ell}} 2L^*(\theta) = n\{- \alpha_n \log|S_1| - \beta_n \log|S_2| - p + \sum_{i=1}^{\ell} \sum_{j\in\kappa_i} (\alpha_n \log \hat{c}_i + \beta_n \log\delta_j - \log(\alpha_n \hat{c}_i + \beta_n \delta_j))\}. \tag{6.12}$$

Also, we have

$$\sup_{\Sigma_1>0,\Sigma_2>0} 2L^*(\theta) = n\{- \alpha_n \log|S_1| - \beta_n \log|S_2| - p\}. \tag{6.13}$$

Now, let

$$L^*(k_1,\ldots,k_\ell) = \sup_{\theta\in\Theta^*_{k_1,\ldots,k_\ell}} \{2L^*(\theta)\} - \sup_{\Sigma_1>0,\Sigma_2>0} \{2L^*(\theta)\}. \tag{6.14}$$

Then, we get

$$L^*(k_1,\ldots,k_\ell) = n\sum_{i=1}^{\ell}\sum_{j\in\kappa_i}(\alpha_n\log\hat{c}_i + \beta_n\log\delta_j - \log(\alpha_n\hat{c}_i + \beta_n\delta_j)) \tag{6.15}$$

Now, let

$$G^*(k_1,\ldots,k_\ell) = L^*(k_1,\ldots,k_\ell) - \ell C_n \tag{6.16}$$

where C_n satisfies the following conditions:

$$\text{(i)}\quad \lim_{n\to\infty}(C_n/n) = 0$$

$$\text{(ii)}\quad \lim_{n\to\infty}(C_n/\log\log n) = \infty \tag{6.17}$$

Then, we estimate $(r,q_1,\ldots,q_r)$ with $(\hat{r},\hat{q}_1,\ldots,\hat{q}_{\hat{r}})$ where

$$G^*(\hat{q}_1,\ldots,\hat{q}_{\hat{r}}) = \max_{(k_1,\ldots,k_\ell)\in K} G^*(k_1,\ldots,k_\ell) \tag{6.18}$$

where $\kappa = \{(k_1,\ldots,k_\ell),\ k_1 < k_2 < \ldots < k_\ell = p,\ \ell = 1,2,\ldots,p\}$.
We now prove the strong consistency of the above procedure.

THEOREM 6.1. Let n_1S_1 and n_2S_2 be distributed independently as central Wishart matrices with n_1 and n_2 degrees of freedom respectively. Also, let $E(S_i) = \Sigma_i$ $(i = 1,2)$. Then $(\hat{r},\hat{q}_1,\ldots,\hat{q}_{\hat{r}})$ defined by (6.16) is a strongly consistent estimate of $(r,q_1,\ldots,q_r)$ when $n\to\infty$, $\alpha_n\in[a,b]\subset(0,1)$ with a and b geing constants and the true model $M^*_{q_1\ldots q_r}$.

PROOF. Suppose that $M^*_{q_1 \ldots q_r}$ is the true model and $\lambda_1 \geq \cdots \geq \lambda_p$ are the eigenvalues of $\Sigma_2 \Sigma_1^{-1}$. By the law of the iterated logarithm, we have

$$S_i - \Sigma_i = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.} \quad \text{as} \quad n \to \infty,$$

for $i = 1,2$. Thus,

$$S_1^{-\frac{1}{2}} S_2 S_1^{-\frac{1}{2}} - \Sigma_1^{-\frac{1}{2}} \Sigma_2 \Sigma_1^{-\frac{1}{2}} = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.} \quad \text{as } n \to \infty$$

By Lemma 3.2 in (Zhao, Krishnaiah and Bai (1985)),

$$|\delta_i - \lambda_i| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.} \quad \text{as } n \to \infty \tag{6.19}$$

for $i = 1,\ldots,p$. Suppose that $(k_1,\ldots,k_\ell) \supseteq (j_1,\ldots,j_m)$. Then

$$\sup_{\theta \in \Theta^*_{k_1,\ldots,k_\ell}} 2L^*(\theta) \geq \sup_{\theta \in \Theta^*_{j_1,\ldots,j_m}} 2L^*(\theta),$$

so, we have

$$L^*(k_1,\ldots,k_\ell) \geq L^*(j_1,\ldots,j_m). \tag{6.20}$$

Now we suppose that $(k_1,\ldots,k_\ell) \supset (q_1,\ldots,q_r)$. Then $\ell > r$. Write $\kappa_i = \{q_{i-1}+1, q_{i-1}+2,\ldots,q_i\}$, $\Delta q_i = a_i$ and $\nu_i = (\delta_i - \lambda_i)/\lambda_i$. Also, put $\hat{\mu}_i = \dfrac{\hat{c}_i - c_i}{c_i}$ for $i = 1,2,\ldots,r$. Assume $(1-\varepsilon_n)\, c_i \leq \delta_j \leq (1+\varepsilon_n) c_i$ for all $j \in \kappa_i$. Then

$$\frac{a_i}{\alpha_n \hat{c}_i c_i^{-1}(1-\varepsilon_n)^{-1} + \beta_n} \leq \sum_{j \in \kappa_i} \frac{\delta_j}{\alpha_n \hat{c}_i + \beta_n \delta_j} \leq \frac{a_i}{\alpha_n \hat{c}_i c_i^{-1}(1+\varepsilon_n)^{-1} + \beta_n} \tag{6.21}$$

Since

$$\sum_{j\in\kappa_i} \frac{\delta_j}{\alpha_n \hat{c}_i + \beta_n \delta_j} = a_i,$$

we know that

$$|\hat{c}_i - c_i| \leq \varepsilon_n c_i \text{ for } i = 1,\ldots,r. \tag{6.22}$$

Using (6.19) and (6.22), we have

$$\hat{\mu}_i = (\hat{c}_i - c_i)/c_i = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ a.s.} \tag{6.23}$$

as $n \to \infty$ for $i = 1,\ldots,r$. By (6.15) and Taylor's expansion, we get

$$0 \leq L^*(k_1,\ldots,k_\ell) - L^*(q_1,\ldots,q_r) \leq - L^*(q_1,\ldots,q_r) \tag{6.24}$$

$$= n \sum_{i=1}^{r} \sum_{j\in\kappa_i} [\log(1+\alpha_n\hat{\mu}_i+\beta_n\nu_j) - \alpha_n\log(1+\hat{\mu}_i) - \beta_n\log(1+\nu_j)]$$

$$\leq \frac{n}{2} \sum_{i=1}^{r} \sum_{j\in\kappa_i} [\alpha_n\hat{\mu}_i^2+\beta_n\nu_j^2 + (\alpha_n\hat{\mu}_i+\beta_n\nu_j)^2](1+o(1)) \text{ a.s.}$$

Now using (6.19) and (6.23), we have

$$0 \leq - L^*(q_1,\ldots,q_r) = O(\log\log n) \text{ a.s.} \tag{6.25}$$

as $n \to \infty$. By (6.16),(6.17), (6.24) and (6.25), with probability one for large n,

$$G^*(q_1,\ldots,q_r) - G^*(k_1,\ldots,k_\ell) = L^*(q_1,\ldots,q_r) - L^*(k_1,\ldots,k_\ell) + (\ell-r)C_n > 0. \tag{6.26}$$

Finally we suppose that $(k_1,\ldots,k_\ell)$ is a partition of $[o,p]$ such that there exists at least one q_t satisfying $k_{i-1} < q_t < k_i$ for some i. Define a new partition $(1,2,\ldots,q_t-1,q_t+1,\ q_t+2,\ldots,p) \stackrel{\Delta}{=} (j_1,\ldots,j_{p-1})$. By the fact $(j_1,\ldots,j_{p-1}) \supseteq (k_1,\ldots,k_\ell)$ and (6.20), we have

$$L^*(q_1,\ldots,q_r) - L^*(k_1,\ldots,k_\ell) \geq L^*(q_1,\ldots,q_r) - L^*(j_1,\ldots,j_{p-1}) \tag{6.27}$$

Now, let $N_t = \{q_t,\ q_{t+1}\}$. It is easy to see that

$$-L^*(j_1,\ldots,j_{p-1}) = n \sum_{j\in N_t} [\log(\alpha_n\hat{\lambda}+\beta_n\delta_j) - \alpha_n\log\hat{\lambda} + \beta_n\log\delta_j] \tag{6.28}$$

where

$$\sum_{j\in N_t} \frac{\delta_j}{\alpha_n\hat{\lambda}+\beta_n\delta_j} = 2. \tag{6.29}$$

Define μ_n such that

$$\sum_{j\in N_t} \frac{\lambda_j}{\alpha_n\mu_n+\beta_n\lambda_j} = 2. \tag{6.30}$$

By $\lim_{n\to\infty} \delta_j = \lambda_j$ a.s. for $j\in N_t$, we have

$$\lim_{n\to\infty}(\hat{\lambda}-\mu_n) = 0 \quad \text{a.s.} \tag{6.31}$$

By (6.31), $c_t > c_{t+1}$, and the condition $\alpha_n\in[a,b]\subset(0,1)$, there exists a constant μ_o such that

$$c_t > \mu_o > \mu_n, \quad n = 1,2,\ldots\ . \tag{6.32}$$

By (6.29), (6.31) and (6.32), we get

$$\begin{aligned} -L^*(j_1,\ldots,j_{p-1}) &\stackrel{\text{a.s.}}{=\!=} n \sum_{j\in N_t} [\log(\alpha_n\mu_n+\beta_n\lambda_j) - \alpha_n\log\mu_n \\ &\qquad - \beta_n\log\lambda_j] + o(n) \\ &\geq n[\log(\alpha_n\mu_n+\beta_n c_t) - \alpha_n\log\mu_n - \beta_n\log c_t] + o(n) \end{aligned}$$

$$\geq n[\log(\alpha_n\mu_o+\beta_n c_t) - \alpha_n\log\mu_o - \beta_n\log c_t] + o(n) \tag{6.33}$$

The function

$$h(\alpha) = \log(\alpha\mu_o+\beta c_t) - \alpha_n\log\mu_o - \beta_n\log c_t, \ (\beta=1-\alpha) \tag{6.34}$$

is positive and continuous for $\alpha\in[a,b]$. So there exists a constant $\alpha_o\in[a,b]$ such that

$$h(\alpha_n) \geq h(\alpha_o) = \log(\alpha_o\mu_o+\beta_o c_t) - \alpha_o\log\mu_o - \beta_o\log c_t > 0. \tag{6.35}$$

By applying (6.33) - (6.35), we get, with probability one

$$-L^*(j_1,\ldots,j_{p-1}) \geq \tfrac{1}{2}nh(\alpha_o) \tag{6.36}$$

for large n. From (6.25),(6.27),(6.36) and $\lim_{n\to\infty} C_n/n = 0$, we know that, with probability one for large n,

$$\begin{aligned} G^*(q_1,\ldots,q_r) - G^*(k_1,\ldots,k_\ell) &= L^*(q_1,\ldots,q_r) - L^*(k_1,\ldots,k_\ell) + o(n) \\ &\geq \frac{1}{2} nh(\alpha_o) + o(n) > 0. \end{aligned} \tag{6.37}$$

From (6.26) and (6.37) it follows that, with probability one for large n,

$$(\hat{r},\hat{q}_1,\ldots,\hat{q}_{\hat{r}}) = (r,q_1,\ldots,q_r), \tag{6.38}$$

which completes the proof of Theorem 6.1.

<u>REMARK 6.1</u>. $L^*(k_1,\ldots,k_\ell)$ can be regarded as a general test statistic (not necessarily LRT statistic) for testing the hypothesis $H^*_{k_1\ldots k_\ell}$. Also, let n_1S_1 and n_2S_2 be distributed as $\sum_{i=1}^{n_1} \underset{\sim}{X}_i\underset{\sim}{X}_i'$ and $\sum_{i=1}^{n_2} \underset{\sim}{Y}_i\underset{\sim}{Y}_i'$ respectively where $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_{n_1}$ and $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_{n_2}$

are subject to the following conditions:

(a) $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_{n_1}$ are i.i.d. and $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_{n_2}$ are i.i.d. such that $E(\underset{\sim}{X}_1)=E(\underset{\sim}{Y}_1)=0$

(b) $(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_{n_1})$ and $(\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_{n_2})$ are independent or not

(c) $E(\underset{\sim}{X}_1\underset{\sim}{X}_1') = \Sigma_1$ and $E(\underset{\sim}{Y}_1\underset{\sim}{Y}_1') = \Sigma_2$ are positive definite

(d) $E(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^2 < \infty$ and $E(\underset{\sim}{Y}_1'\underset{\sim}{Y}_1)^2 < \infty$

Then, $(\hat{r},\hat{q}_1,\ldots,\hat{q}_{\hat{r}})$ is still a strongly consistent estimate of $(r,q_1,\ldots,q_r)$.

<u>REMARK 6.2.</u> If $\lambda_p = c_\ell$ is known, we can assume $c_\ell = 1$. In this case, we redenote $H^*_{k_1\ldots k_\ell}$, $\Theta^*_{k_1,\ldots,k_\ell}$, $M^*_{k_1,\ldots,k_\ell}$, $L^*_{k_1,\ldots,k_\ell}$ and $G^*(k_1,\ldots,k_\ell)$ by the corresponding notations without stars. Following the same lines as in the case of proof from (6.1) to (6.10), we find that

$$\sup_{\Theta(k_1\ldots k_\ell)} 2L^*(\theta) = \sup_{c_1>\ldots>c_{\ell-1}>1} n\{-\log|\hat{R}\hat{R}| + \alpha_n \sum_{i=1}^{p} \log\lambda_i - \sum_{i=1}^{p} \log(\alpha_n\lambda_i+\beta_n\delta_i)p\}$$

$$= \sup_{c_1>\ldots>c_{\ell-1}>1} n\{-\log|\hat{R}\hat{R}| - p + \alpha_n \sum_{i=1}^{\ell-1} \Delta k_i \log c_i - \sum_{i=1}^{\ell-1} \sum_{j\in K_i} \log(\alpha_n c_i+\beta_n\delta_i)$$

$$- \sum_{j\in K_\ell} \log(\alpha_n+\beta_n\delta_i)\} \tag{6.39}$$

Define $\hat{c}_i$ as those defined in (6.11) and define

$$\tau = \max\{i \leq \ell-1,\ \hat{c}_i > 1\} \tag{6.40}$$

If $i \leq \tau$, then

$$\sup_{c_i>1} (\alpha_n \Delta k_i \log c_i - \sum_{j\in K_i} \log(\alpha_n c_i+\beta_n\delta_i)) =$$

$$= \alpha_n \Delta k_i \log\hat{c}_i - \sum_{j\in K_i} \log(\alpha_n c_i+\beta_n\delta_i) \tag{6.41}$$

and the superium can be reached at $c_i = \hat{c}_i$. If $i > \tau$, i.e., $\hat{c}_i \leq 1$, for $c_i > 1$ we have

$$\frac{d}{dc_i}(\alpha_n \Delta k_i \log c_i - \sum_{j\in\kappa_i} \log(\alpha_n c_i + \beta_n \delta_i))$$

$$= \frac{\alpha_n \Delta k_i}{c_i} - \sum_{j\in\kappa_i} \frac{\alpha_n}{\alpha_n c_i + \beta_n \delta_i} = \frac{\alpha_n}{c_i}\left(\Delta k_i - \sum_{j\in\kappa_i} \frac{c_i}{\alpha_n c_i + \beta_n \delta_i}\right)$$

$$< \frac{\alpha_n}{c_i}\left(\Delta k_i - \sum_{j\in\kappa_i} \frac{\hat{c}_i}{\alpha_n \hat{c}_i + \beta_n \delta_i}\right) = 0,$$

Hence

$$\operatorname*{Sup}_{c_i > 1} (\alpha_n \Delta k_i \log c_i - \sum_{j\in\kappa_i} \log(\alpha_n c_i + \beta_n \delta_i)) = - \sum_{j\in\kappa_i} \log(\alpha_n + \beta_n \delta_i),$$

and the superium is reached at $c_i = 1$. Noting that $\hat{c}_1 > \ldots > \hat{c}_\tau > 1$, we obtain

$$\operatorname*{Sup}_{\Theta(k_1 \ldots k_\ell)} 2L^*(\theta) = n\{- \alpha_n \log|S_1| - \beta_n \log|S_2| - p$$

$$+ \sum_{i=1}^{\tau} \sum_{j\in\kappa_i} (\alpha_n \log\hat{c}_i + \beta_n \log\delta_j - \log(\alpha_n \hat{c}_i + \beta_n \delta_i)] - \sum_{i=\tau+1}^{\ell} \log(\alpha_n + \beta_n \delta_i)\} \tag{6.42}$$

Hence we get

$$L(k_1 \ldots k_\ell) = n \sum_{i=1}^{\tau} \sum_{j\in\kappa_i} [\alpha_n \log\hat{c}_i + \beta_n \log\delta_i - \log(\alpha_n \hat{c}_i + \beta_n \delta_i)]$$

$$- n \sum_{i=\tau+1}^{\ell} \log(\alpha_n + \beta_n \delta_i) \tag{6.43}$$

Now, let

$$G(k_1,\ldots,k_\ell) = L(k_1,\ldots,k_\ell) - \ell c_n \tag{6.44}$$

where c_n satisfies (6.17). Then we estimate $(r, q_1, \ldots, q_r)$ with $(\hat{r}, \hat{q}_1, \ldots, \hat{q}_{\hat{r}})$, where

$$G(\hat{q}_1,\ldots,\hat{q}_{\hat{r}}) = \max_{(k_1,\ldots,k_\ell)\in K} G(k_1,\ldots,k_\ell), \tag{6.45}$$

where K is the set defined below (6.18). Similarily as Theorem 6.1, we have

THEOREM 6.2. Let n_1S_2, n_2S_2 be distributed independently as central Wishart matrices with n_1 and n_2 degrees of freedom respectively. Also, let $ES_i = \Sigma_i (i=1,2)$. Then $(\hat{r},\hat{q}_1,\ldots,\hat{q}_{\hat{r}})$ defined by (6.45) is a strongly consistent estimate of $(r,q_1,\ldots,q_r)$, when $n \to \infty$, $\alpha_n \epsilon [a,b] \subset (0,1)$ with a,b being constants and the true model $M_{q_1\ldots q_r}$.

Since, with probability one, when large enough, $\tau = r - 1$ under the hypothesis $H_{q_1\ldots q_r}$, we find that Theorem 6.2 can be proved by the same argument as used in the proof of Theorem 6.1. So that we omit it here.

REMARK 6.3. Remark 6.1 concerning Theorem 6.1 is also available to Theorem 6.2.

REMARK 6.4. If n_1S_1 and n_2S_2 are distributed independently as central complex Wishart matrices. Then the log-likelihood function $L(\theta)$ is given by, up to an adding constant,

$$L(\theta) = - n_1\log|\Sigma_1| - n_2\log|\Sigma_2| - n_1 tr\Sigma_1^{-1}S_1 - n_2 tr\Sigma_2^{-1}S_2,$$

In the arguments in Sections 3-6, we only need change the following notations

$$\begin{aligned} \Sigma_1 &= RR^*, & \Sigma_2 &= R\Lambda R^* \\ S_1 &= \hat{R}\hat{R}^* & S_2 &= \hat{R}\Delta\hat{R}^* \quad \text{in (3.2)} \end{aligned}$$

where A^* denotes the transpose of the conjugate of the matrix A.

$$L_1(V',\Lambda) = (n_1+n_2)\log|V^*V| - n_1 trV^*V - n_2 tr\Lambda^{-1}V\Delta V^* \quad \text{in (3.4)}$$

and Q being a Hermite matrix instead of an orthogonal matrix and rewrite Q' as Q^* in (3.7) and (3.8). Finally we can get the same representations of the log-likelihood ratio test statistic as given in (3.18), (3.23),(6.15) and (6.43). By the same way, we can prove the analogues to Theorem 4.1, 6.1 and 6.2.

REFERENCES

[1] AKAIKE, H. (1972). Information theory and an extension of the maximum likelihood principle. In Proceedings of the Second International Symposium on Information Theory, Supp. to Problems of Control and Information Theory, pp. 267-281.

[2] ANDERSON, T.W. (1984). Estimating linear statistical relationships. Ann. Statist., **12**, 1-45.

[3] ANDERSON, T.W. (1985). Components of variance in MANOVA. In Multivariate Analysis VI (P.R. Krishnaiah, editor), pp 1-8 North-Holland Publishing Co.

[4] FANG, C and KRISHNAIAH, P.R. (1981). Asymptotic joint distributions of functions of the eigenvalues of some random matrices for nonnormal populations. J. Multivariate Anal., **12**, 39-63.

[5] HANNAN, E.J. and QUINN, B.G. (1979). The determination of the order of an autoregression. J. Roy. Statist. Soc. Ser. B, **41**, 190-195.

[6] KRISHNAIAH, P.R. (1976). Some recent developments on complex multivariate distributions. J. Multivariate Anal., **6**, 1-30.

[7] KRISHNAIAH, P.R. and SCHUURMANN, F.J. Computations of Complex Multivariate Distributions (in preparation). To be published by North-Holland Publishing Co.

[8] KRISHNAIAH, P.R. (1980). Computations of some multivariate distributions. In Handbook of Statistics, Vol. 1: Analysis of Variance. (P.R.Krishnaiah, editor). North-Holland, Amsterdam.

[9] KUMARESAN, R. and TUFTS, D.W. (1980). Data adaptive principal component signal processing. Proc. 19th IEEE Conf. Decision and Control, Albuquerque, NM, pp. 949-954.

[10] LIGGETT, W.S. (1973). Passive sonar: fitting models to multiple time series. In Signal Processing (J.W. Griffith et al, editors) pp. 327-345. Academic Press, New York.

[11] MURIHEAD, R.J. (1978). Latent roots and matrix variates: a review of some asymptotic results. Ann. Statist., **6**, 5-33.

[12] RAO, C.R.(1983). Likelihood ratio tests for relationships between two covariance matrices. In Studies in Econometrics, Time Series and Multivariate Statistics (T. Amemiya, S. Karlin and L. Goodman, editors). Academic Press, Inc. New York.

[13] RISSANEN, J. (1978). Modeling by shortest data description. Automatica, **14**, 465-471.

[14] SCHMIDT, R.O. (1979). Multiple emitter location and signal parameter estimation. In Proc. RADC Spectrum Estimation Workshop, Rome.

[15] SCHOTT, J.R. and SAW, J.G. (1984). A multivariate one-way classification model with random effects. J. Multivariate Anal. 15, 1-12.

[16] SCHWARTZ, G. (1978). Estimating the dimension of a model. Ann. Statist., 6, 461-464.

[17] TUFTS, D. KIRSTEINS, I. and KUMARESAN, R. (1983). Data-adaptive detection of a weak signal. IEEE Transactions on Aerospace and Electronic Systems. Vol. AES-19, 313-316.

[18] WAX. M. and KAILATH, T. (1984). Determination of the number of signals by information theoretic criteria. IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-33.

[19] WAX, M. SHAN, T.J., and KAILATH, T. (1984). Spatio-temporal spectral analysis by eigenstructure methods. IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-32, 817-827.

[20] WOODING, R. A. (1956). The multivariate distribution of complex normal variables. Biometrika, 43, 212-215.

[21] ZHAO, L.C., KRISHNAIAH, P.R. and BAI, Z.D. (1985). On detection of number of signals in presence of white noise. Tech. Rept. No. 85-37. Center for Multivariate Analysis, University of Pittsburgh.